



4

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Scaglioni, Pier Paolo
Melegari, Margherita
Wands, Jack
- (ii) TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
- (iii) NUMBER OF SEQUENCES: 14
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Fish & Richardson P.C.
 - (B) STREET: 225 Franklin Street
 - (C) CITY: Boston
 - (D) STATE: MA
 - (E) COUNTRY: USA
 - (F) ZIP: 02110-2804
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/667,073
 - (B) FILING DATE: 20-JUN-1996
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/492,489
 - (B) FILING DATE: 20-JUN-1995
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Fraser, Janis K. (formerly Clark, Paul T.)
 - (B) REGISTRATION NUMBER: 34,819 (formerly 30,162)
 - (C) REFERENCE/DOCKET NUMBER: 10978/282001 (formerly 00786/282001)
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 617/542-5070
 - (B) TELEFAX: 617/542-8906
 - (C) TELEX: 200154

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1041 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG GAC ATA GAT CCC TAT AAA GAA TTT GGT TCA TCT TAT CAG TTG TTG
Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ser Ser Tyr Gln Leu Leu
1 5 10 15

48

AAT TTT CTT CCT TTG GAC TTC TTT CCT GAC CTT AAT GCT TTG GTG GAC	96
Asn Phe Leu Pro Leu Asp Phe Phe Pro Asp Leu Asn Ala Leu Val Asp	
20 25 30	
ACT GCT ACT GCC TTG TAT GAA GAA GAG CTA ACA GGT AGG GAA CAT TGC	144
Thr Ala Thr Ala Leu Tyr Glu Glu Glu Leu Thr Gly Arg Glu His Cys	
35 40 45	
TCT CCG CAC CAT ACA GCT ATT AGA CAA GCT TTA GTA TGC TGG GAT GAA	192
Ser Pro His His Thr Ala Ile Arg Gln Ala Leu Val Cys Trp Asp Glu	
50 55 60	
TTA ACT AAA TTG ATA GCT TGG ATG AGC TCT AAC ATA ACT TCT GAA CAA	240
Leu Thr Lys Leu Ile Ala Trp Met Ser Ser Asn Ile Thr Ser Glu Gln	
65 70 75 80	
GTA AGA ACA ATC ATA GTA AAT CAT GTC AAT GAT ACC TGG GGA CTT AAG	288
Val Arg Thr Ile Ile Val Asn His Val Asn Asp Thr Trp Gly Leu Lys	
85 90 95	
GTG AGA CAA AGT TTA TGG TTT CAT TTG TCA TGT CTC ACT TTC GGA CAA	336
Val Arg Gln Ser Leu Trp Phe His Leu Ser Cys Leu Thr Phe Gly Gln	
100 105 110	
CAT ACA GTT CAA GAA TTT TTA GTA AGT TTT GTA GTA TGG ATC AGA ACT	384
His Thr Val Gln Glu Phe Leu Val Ser Phe Val Val Trp Ile Arg Thr	
115 120 125	
CCA GCT CCA TAT AGA CCT CCT AAT GCA CCC ATT CTC TCG ACT CTT CCG	432
Pro Ala Pro Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro	
130 135 140	
GAA CAT ACA GTC ATT AGA AGA GGA GGT GCA AGA GCT TCT AGG TCC CCC	480
Glu His Thr Val Ile Arg Arg Gly Gly Ala Arg Ala Ser Arg Ser Pro	
145 150 155 160	
AGA AGA CGC ACT CCC TCT CCT CGC AGG AGA AGA TCC CAA AAT TCG CAG	528
Arg Arg Arg Thr Pro Ser Pro Arg Arg Arg Ser Ser Gln Asn Ser Gln	
165 170 175	
TTC CAA ACT TGC AAA CAC TTG CCA ACC TCC TGT CCA CCA ACT TGC AAT	576
Phe Gln Thr Cys Lys His Leu Pro Thr Ser Cys Pro Pro Thr Cys Asn	
180 185 190	
GGC TTT CGT TGG ATG TAT CTG CGG CGT TTT ATC ATA TAC CTA TTA GTC	624
Gly Phe Arg Trp Met Tyr Leu Arg Arg Phe Ile Ile Tyr Leu Leu Val	
195 200 205	
CTG CTG CTG TGC CTC ATC TTC TTG TTG GTT CTC CTG GAC TGG AAA GGT	672
Leu Leu Leu Cys Leu Ile Phe Leu Leu Val Leu Leu Asp Trp Lys Gly	
210 215 220	
TTA ATA CCT GTC TGT CCT CTT CAA CCC ACA ACA GAA ACA ACA GTC AAT	720
Leu Ile Pro Val Cys Pro Leu Gln Pro Thr Thr Glu Thr Thr Val Asn	
225 230 235 240	
TGC AGA CAA TGC ACA ATC TCT GCA CAA GAC ATG TAT ACT CCT CCT TAC	768
Cys Arg Gln Cys Thr Ile Ser Ala Gln Asp Met Tyr Thr Pro Pro Tyr	
245 250 255	

TGT TGT TGT TTA AAA CCT ACG GCA GGA AAT TGC ACT TGT TGG CCC ATC	816
Cys Cys Cys Leu Lys Pro Thr Ala Gly Asn Cys Thr Cys Trp Pro Ile	
260 265 270	
CCT TCA TCA TGG GCT TTA GGA AAT TAC CTA TGG GAG TGG GCC TTA GCT	864
Pro Ser Ser Trp Ala Leu Gly Asn Tyr Leu Trp Glu Trp Ala Leu Ala	
275 280 285	
CGT CTC TCT TGG CTC AAT TTA CTA GTG CCC TTG CTT CAA TGG TTA GGA	912
Arg Leu Ser Trp Leu Asn Leu Leu Val Pro Leu Leu Gln Trp Leu Gly	
290 295 300	
GGA ATT TCC CTC ATT GCG TGG TTT TTG CTT ATA TGG ATG ATT TGG TTT	960
Gly Ile Ser Leu Ile Ala Trp Phe Leu Leu Ile Trp Met Ile Trp Phe	
305 310 315 320	
TGG GGG CCC GCA CTT CTG AGC ATC TTA CCG CCA TTT ATT CCC ATA TTT	1008
Trp Gly Pro Ala Leu Leu Ser Ile Leu Pro Pro Phe Ile Pro Ile Phe	
325 330 335	
GTT CTG TTT TTC TTG ATT TGG GTA TAC ATT TGA	1041
Val Leu Phe Phe Leu Ile Trp Val Tyr Ile	
340 345	

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 346 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ser Ser Tyr Gln Leu Leu	
1 5 10 15	
Asn Phe Leu Pro Leu Asp Phe Phe Pro Asp Leu Asn Ala Leu Val Asp	
20 25 30	
Thr Ala Thr Ala Leu Tyr Glu Glu Glu Leu Thr Gly Arg Glu His Cys	
35 40 45	
Ser Pro His His Thr Ala Ile Arg Gln Ala Leu Val Cys Trp Asp Glu	
50 55 60	
Leu Thr Lys Leu Ile Ala Trp Met Ser Ser Asn Ile Thr Ser Glu Gln	
65 70 75 80	
Val Arg Thr Ile Ile Val Asn His Val Asn Asp Thr Trp Gly Leu Lys	
85 90 95	
Val Arg Gln Ser Leu Trp Phe His Leu Ser Cys Leu Thr Phe Gly Gln	
100 105 110	
His Thr Val Gln Glu Phe Leu Val Ser Phe Val Val Trp Ile Arg Thr	
115 120 125	
Pro Ala Pro Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro	
130 135 140	
Glu His Thr Val Ile Arg Arg Gly Gly Ala Arg Ala Ser Arg Ser Pro	
145 150 155 160	
Arg Arg Arg Thr Pro Ser Pro Arg Arg Arg Ser Gln Asn Ser Gln	
165 170 175	
Phe Gln Thr Cys Lys His Leu Pro Thr Ser Cys Pro Pro Thr Cys Asn	
180 185 190	
Gly Phe Arg Trp Met Tyr Leu Arg Arg Phe Ile Ile Tyr Leu Leu Val	
195 200 205	

Leu	Leu	Leu	Cys	Leu	Ile	Phe	Leu	Leu	Val	Leu	Leu	Asp	Trp	Lys	Gly
210						215					220				
Leu	Ile	Pro	Val	Cys	Pro	Leu	Gln	Pro	Thr	Thr	Glu	Thr	Thr	Val	Asn
225					230					235					240
Cys	Arg	Gln	Cys	Thr	Ile	Ser	Ala	Gln	Asp	Met	Tyr	Thr	Pro	Pro	Tyr
				245					250					255	
Cys	Cys	Cys	Leu	Lys	Pro	Thr	Ala	Gly	Asn	Cys	Thr	Cys	Trp	Pro	Ile
			260					265					270		
Pro	Ser	Ser	Trp	Ala	Leu	Gly	Asn	Tyr	Leu	Trp	Glu	Trp	Ala	Leu	Ala
		275				280						285			
Arg	Leu	Ser	Trp	Leu	Asn	Leu	Leu	Val	Pro	Leu	Leu	Gln	Trp	Leu	Gly
	290				295					300					
Gly	Ile	Ser	Leu	Ile	Ala	Trp	Phe	Leu	Leu	Ile	Trp	Met	Ile	Trp	Phe
305					310					315					320
Trp	Gly	Pro	Ala	Leu	Ser	Ile	Leu	Pro	Pro	Phe	Ile	Pro	Ile	Phe	
			325					330					335		
Val	Leu	Phe	Phe	Leu	Ile	Trp	Val	Tyr	Ile						
		340					345								

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1056 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCT	CGC	AGG	ATG	GAC	ATC	GAC	CCT	TAT	AAA	GAA	TTT	GGA	GCT	ACT	GTG	48
Pro	Arg	Arg	Met	Asp	Ile	Asp	Pro	Tyr	Lys	Glu	Phe	Gly	Ala	Thr	Val	
1				5				10						15		
GAG	TTA	CTC	TCG	TTT	TTG	CCT	TCT	GAC	TTC	TTT	CCT	TCA	GTA	CGA	GAT	96
Glu	Leu	Leu	Ser	Phe	Leu	Pro	Ser	Asp	Phe	Phe	Pro	Ser	Val	Arg	Asp	
			20					25					30			
CTT	CTA	GAT	ACC	GCC	TCA	GCT	CTG	TAT	CGG	GAA	GCC	TTA	GAG	TCT	CCT	144
Leu	Leu	Asp	Thr	Ala	Ser	Ala	Leu	Tyr	Arg	Glu	Ala	Leu	Glu	Ser	Pro	
		35					40					45				
GAG	CAT	TGT	TCA	CCT	CAC	CAT	ACT	GCA	CTC	AGG	CAA	GCA	ATT	CTT	TGC	192
Glu	His	Cys	Ser	Pro	His	His	Thr	Ala	Leu	Arg	Gln	Ala	Ile	Leu	Cys	
	50					55				60						
TGG	GGG	GAA	CTA	ATG	ACT	CTA	GCT	ACC	TGG	GTG	GGT	GTT	AAT	TTG	GAA	240
Trp	Gly	Glu	Leu	Met	Thr	Leu	Ala	Thr	Trp	Val	Gly	Val	Asn	Leu	Glu	
65				70				75					80			
GAT	CCA	GCG	TCT	AGA	GAC	CTA	GTA	GTC	AGT	TAT	GTC	AAC	ACT	AAT	ATG	288
Asp	Pro	Ala	Ser	Arg	Asp	Leu	Val	Val	Ser	Tyr	Val	Asn	Thr	Asn	Met	
			85					90					95			
GGC	CTA	AAG	TTC	AGG	CAA	CTC	TTG	TGG	TTT	CAC	ATT	TCT	TGT	CTC	ACT	336
Gly	Leu	Lys	Phe	Arg	Gln	Leu	Leu	Trp	Phe	His	Ile	Ser	Cys	Leu	Thr	
		100					105						110			
TTT	GGA	AGA	GAA	ACA	GTT	ATA	GAG	TAT	TTG	GTG	TCT	TTC	GGA	GTG	TGG	384
Phe	Gly	Arg	Glu	Thr	Val	Ile	Glu	Tyr	Leu	Val	Ser	Phe	Gly	Val	Trp	
		115					120					125				

ATT CGC ACT CCT CCA GCT TAT AGA CCA CCA AAT GCC CCT ATC CTA TCA Ile Arg Thr Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser 130 135 140	432
ACA CTT CCG GAA CAT ACA GTC ATT AGA AGA GGA GGT GCA AGA GCT TCT Thr Leu Pro Glu His Thr Val Ile Arg Arg Gly Gly Ala Arg Ala Ser 145 150 155 160	480
AGG TCC CCC AGA AGA CGC ACT CCC TCT CCT CGC AGG AGA AGA TCC CAA Arg Ser Pro Arg Arg Arg Thr Pro Ser Pro Arg Arg Arg Ser Gln 165 170 175	528
AAT TCG CAG TCC CCA ACC TCC AAT CAC TCA CCA ACC TCT TGT CCT CCA Asn Ser Gln Ser Pro Thr Ser Asn His Ser Pro Thr Ser Cys Pro Pro 180 185 190	576
ACT TGT CCT GGT TAT CGC TGG ATG TGT CTG CGG CGT TTT ATC ATC TTC Thr Cys Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg Phe Ile Ile Phe 195 200 205	624
CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC TTG TTG GTT CTT CTG GAC Leu Phe Ile Leu Leu Leu Cys Leu Ile Phe Leu Leu Val Leu Leu Asp 210 215 220	672
TAT CAA GGT ATG TTG CCC GTT TGT CCT CTA ATT CCA GGA TCC TCA ACA Tyr Gln Gly Met Leu Pro Val Cys Pro Leu Ile Pro Gly Ser Ser Thr 225 230 235 240	720
ACC AGC ACG GGA CCA TGC CGG ACC TGC ATG ACT ACT GCT CAA GGA ACC Thr Ser Thr Gly Pro Cys Arg Thr Cys Met Thr Thr Ala Gln Gly Thr 245 250 255	768
TCT ATG TAT CCC TCC TGT TGC TGT ACC AAA CCT TCG GAC GGA AAT TGC Ser Met Tyr Pro Ser Cys Cys Cys Thr Lys Pro Ser Asp Gly Asn Cys 260 265 270	816
ACC TGT ATT CCC ATC CCA TCA TCC TGG GCT TTC GGA AAA TTC CTA TGG Thr Cys Ile Pro Ile Pro Ser Ser Trp Ala Phe Gly Lys Phe Leu Trp 275 280 285	864
GAG TGG GCC TCA GCC CGT TTC TCC TGG CTC AGT TTA CTA GTG CCA TTT Glu Trp Ala Ser Ala Arg Phe Ser Trp Leu Ser Leu Leu Val Pro Phe 290 295 300	912
GTT CAG TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT TCA GTT ATA Val Gln Trp Phe Val Gly Leu Ser Pro Thr Val Trp Leu Ser Val Ile 305 310 315 320	960
TGG ATG ATG TGG TAT TGG GGG CCA AGT CTG TAC AGC ATC TTG AGT CCC Trp Met Met Trp Tyr Trp Gly Pro Ser Leu Tyr Ser Ile Leu Ser Pro 325 330 335	1008
TTT TTA CCG CTG TTA CCA ATT TTC TTT TGT CTT TGG GTA TAC ATT TAA Phe Leu Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val Tyr Ile 340 345 350	1056

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 351 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Pro	Arg	Arg	Met	Asp	Ile	Asp	Pro	Tyr	Lys	Glu	Phe	Gly	Ala	Thr	Val
1				5					10					15	
Glu	Leu	Leu	Ser	Phe	Leu	Pro	Ser	Asp	Phe	Phe	Pro	Ser	Val	Arg	Asp
			20					25					30		
Leu	Leu	Asp	Thr	Ala	Ser	Ala	Leu	Tyr	Arg	Glu	Ala	Leu	Glu	Ser	Pro
		35					40					45			
Glu	His	Cys	Ser	Pro	His	His	Thr	Ala	Leu	Arg	Gln	Ala	Ile	Leu	Cys
	50					55					60				
Trp	Gly	Glu	Leu	Met	Thr	Leu	Ala	Thr	Trp	Val	Gly	Val	Asn	Leu	Glu
65					70					75				80	
Asp	Pro	Ala	Ser	Arg	Asp	Leu	Val	Val	Ser	Tyr	Val	Asn	Thr	Asn	Met
			85						90					95	
Gly	Leu	Lys	Phe	Arg	Gln	Leu	Leu	Trp	Phe	His	Ile	Ser	Cys	Leu	Thr
			100					105					110		
Phe	Gly	Arg	Glu	Thr	Val	Ile	Glu	Tyr	Leu	Val	Ser	Phe	Gly	Val	Trp
		115					120					125			
Ile	Arg	Thr	Pro	Pro	Ala	Tyr	Arg	Pro	Pro	Asn	Ala	Pro	Ile	Leu	Ser
	130					135					140				
Thr	Leu	Pro	Glu	His	Thr	Val	Ile	Arg	Arg	Gly	Gly	Ala	Arg	Ala	Ser
145					150					155				160	
Arg	Ser	Pro	Arg	Arg	Arg	Thr	Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln
			165						170					175	
Asn	Ser	Gln	Ser	Pro	Thr	Ser	Asn	His	Ser	Pro	Thr	Ser	Cys	Pro	Pro
		180					185						190		
Thr	Cys	Pro	Gly	Tyr	Arg	Trp	Met	Cys	Leu	Arg	Arg	Phe	Ile	Ile	Phe
	195					200						205			
Leu	Phe	Ile	Leu	Leu	Leu	Cys	Leu	Ile	Phe	Leu	Leu	Val	Leu	Leu	Asp
	210				215					220					
Tyr	Gln	Gly	Met	Leu	Pro	Val	Cys	Pro	Leu	Ile	Pro	Gly	Ser	Ser	Thr
225				230					235					240	
Thr	Ser	Thr	Gly	Pro	Cys	Arg	Thr	Cys	Met	Thr	Thr	Ala	Gln	Gly	Thr
			245					250						255	
Ser	Met	Tyr	Pro	Ser	Cys	Cys	Cys	Thr	Lys	Pro	Ser	Asp	Gly	Asn	Cys
		260						265					270		
Thr	Cys	Ile	Pro	Ile	Pro	Ser	Ser	Trp	Ala	Phe	Gly	Lys	Phe	Leu	Trp
	275					280						285			
Glu	Trp	Ala	Ser	Ala	Arg	Phe	Ser	Trp	Leu	Ser	Leu	Leu	Val	Pro	Phe
	290				295						300				
Val	Gln	Trp	Phe	Val	Gly	Leu	Ser	Pro	Thr	Val	Trp	Leu	Ser	Val	Ile
305				310						315				320	
Trp	Met	Met	Trp	Tyr	Trp	Gly	Pro	Ser	Leu	Tyr	Ser	Ile	Leu	Ser	Pro
			325						330				335		
Phe	Leu	Pro	Leu	Leu	Pro	Ile	Phe	Phe	Cys	Leu	Trp	Val	Tyr	Ile	
		340					345						350		

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1194 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATG	GAC	ATC	GAC	CCT	TAT	AAA	GAA	TTT	GGA	GCT	ACT	GTG	GAG	TTA	CTC	48
Met	Asp	Ile	Asp	Pro	Tyr	Lys	Glu	Phe	Gly	Ala	Thr	Val	Glu	Leu	Leu	
1				5					10					15		
TCG	TTT	TTG	CCT	TCT	GAC	TTC	TTT	CCT	TCA	GTA	CGA	GAT	CTT	CTA	GAT	96
Ser	Phe	Leu	Pro	Ser	Asp	Phe	Phe	Pro	Ser	Val	Arg	Asp	Leu	Leu	Asp	
			20					25					30			
ACC	GCC	TCA	GCT	CTG	TAT	CGG	GAA	GCC	TTA	GAG	TCT	CCT	GAG	CAT	TGT	144
Thr	Ala	Ser	Ala	Leu	Tyr	Arg	Glu	Ala	Leu	Glu	Ser	Pro	Glu	His	Cys	
		35					40					45				
TCA	CCT	CAC	CAT	ACT	GCA	CTC	AGG	CAA	GCA	ATT	CTT	TGC	TGG	GGG	GAA	192
Ser	Pro	His	His	Thr	Ala	Leu	Arg	Gln	Ala	Ile	Leu	Cys	Trp	Gly	Glu	
	50					55					60					
CTA	ATG	ACT	CTA	GCT	ACC	TGG	GTG	GGT	GTT	AAT	TTG	GAA	GAT	CCA	GCG	240
Leu	Met	Thr	Leu	Ala	Thr	Trp	Val	Gly	Val	Asn	Leu	Glu	Asp	Pro	Ala	
65					70				75						80	
TCT	AGA	GAC	CTA	GTA	GTC	AGT	TAT	GTC	AAC	ACT	AAT	ATG	GGC	CTA	AAG	288
Ser	Arg	Asp	Leu	Val	Val	Ser	Tyr	Val	Asn	Thr	Asn	Met	Gly	Leu	Lys	
				85					90					95		
TTC	AGG	CAA	CTC	TTG	TGG	TTT	CAC	ATT	TCT	TGT	CTC	ACT	TTT	GGA	AGA	336
Phe	Arg	Gln	Leu	Leu	Trp	Phe	His	Ile	Ser	Cys	Leu	Thr	Phe	Gly	Arg	
			100					105					110			
GAA	ACA	GTT	ATA	GAG	TAT	TTG	GTG	TCT	TTC	GGA	GTG	TGG	ATT	CGC	ACT	384
Glu	Thr	Val	Ile	Glu	Tyr	Leu	Val	Ser	Phe	Gly	Val	Trp	Ile	Arg	Thr	
		115					120					125				
CCT	CCA	GCT	TAT	AGA	CCA	CCA	AAT	GCC	CCT	ATC	CTA	TCA	ACA	CTT	CCG	432
Pro	Pro	Ala	Tyr	Arg	Pro	Pro	Asn	Ala	Pro	Ile	Leu	Ser	Thr	Leu	Pro	
	130					135					140					
GAG	ACT	ACT	GTT	GTT	AGA	CGA	CGA	GGC	AGG	TCC	CCT	AGA	AGA	AGA	ACT	480
Glu	Thr	Thr	Val	Val	Arg	Arg	Arg	Gly	Arg	Ser	Pro	Arg	Arg	Arg	Thr	
145					150				155						160	
CCC	TCG	CCT	CGC	AGA	CGA	AGG	TCT	CAA	TCG	CCG	CGT	CGC	AGA	AGA	TCT	528
Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Pro	Arg	Arg	Arg	Arg	Ser	
			165						170					175		
CAA	TCT	CGG	CTA	GGA	CCC	CTT	CTC	GTG	TTA	CAG	GCG	GGG	TTT	TTC	TTG	576
Gln	Ser	Arg	Leu	Gly	Pro	Leu	Leu	Val	Leu	Gln	Ala	Gly	Phe	Phe	Leu	
			180					185					190			
TTG	ACA	AGA	ATC	CTC	ACA	ATA	CCG	CAG	AGT	CTA	GAC	TCG	TGG	TGG	ACT	624
Leu	Thr	Arg	Ile	Leu	Thr	Ile	Pro	Gln	Ser	Leu	Asp	Ser	Trp	Trp	Thr	
		195					200					205				
TCT	CTC	AAT	TTT	CTA	GGG	GGA	ACT	ACC	GTG	TGT	CTT	GGC	CAA	AAT	TCG	672
Ser	Leu	Asn	Phe	Leu	Gly	Gly	Thr	Thr	Val	Cys	Leu	Gly	Gln	Asn	Ser	
	210					215					220					
CAG	TCC	CCA	ACC	TCC	AAT	CAC	TCA	CCA	ACC	TCT	TGT	CCT	CCA	ACT	TGT	720
Gln	Ser	Pro	Thr	Ser	Asn	His	Ser	Pro	Thr	Ser	Cys	Pro	Pro	Thr	Cys	
225					230					235					240	

CCT GGT TAT CGC TGG ATG TGT CTG CGG CGT TTT ATC ATC TTC CTC TTC	768
Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg Phe Ile Ile Phe Leu Phe	
245 250 255	
ATC CTG CTG CTA TGC CTC ATC TTC TTG TTG GTT CTT CTG GAC TAT CAA	816
Ile Leu Leu Leu Cys Leu Ile Phe Leu Leu Val Leu Leu Asp Tyr Gln	
260 265 270	
GGT ATG TTG CCC GTT TGT CCT CTA ATT CCA GGA TCC TCA ACA ACC AGC	864
Gly Met Leu Pro Val Cys Pro Leu Ile Pro Gly Ser Ser Thr Thr Ser	
275 280 285	
ACG GGA CCA TGC CGG ACC TGC ATG ACT ACT GCT CAA GGA ACC TCT ATG	912
Thr Gly Pro Cys Arg Thr Cys Met Thr Thr Ala Gln Gly Thr Ser Met	
290 295 300	
TAT CCC TCC TGT TGC TGT ACC AAA CCT TCG GAC GGA AAT TGC ACC TGT	960
Tyr Pro Ser Cys Cys Cys Thr Lys Pro Ser Asp Gly Asn Cys Thr Cys	
305 310 315 320	
ATT CCC ATC CCA TCA TCC TGG GCT TTC GGA AAA TTC CTA TGG GAG TGG	1008
Ile Pro Ile Pro Ser Ser Trp Ala Phe Gly Lys Phe Leu Trp Glu Trp	
325 330 335	
GCC TCA GCC CGT TTC TCC TGG CTC AGT TTA CTA GTG CCA TTT GTT CAG	1056
Ala Ser Ala Arg Phe Ser Trp Leu Ser Leu Leu Val Pro Phe Val Gln	
340 345 350	
TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT TCA GTT ATA TGG ATG	1104
Trp Phe Val Gly Leu Ser Pro Thr Val Trp Leu Ser Val Ile Trp Met	
355 360 365	
ATG TGG TAT TGG GGG CCA AGT CTG TAC AGC ATC TTG AGT CCC TTT TTA	1152
Met Trp Tyr Trp Gly Pro Ser Leu Tyr Ser Ile Leu Ser Pro Phe Leu	
370 375 380	
CCG CTG TTA CCA ATT TTC TTT TGT CTT TGG GTA TAC ATT TAA	1194
Pro Leu Leu Pro Ile Phe Cys Leu Trp Val Tyr Ile	
385 390 395	

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 397 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Asp	Ile	Asp	Pro	Tyr	Lys	Glu	Phe	Gly	Ala	Thr	Val	Glu	Leu	Leu
1				5					10					15	
Ser	Phe	Leu	Pro	Ser	Asp	Phe	Phe	Pro	Ser	Val	Arg	Asp	Leu	Leu	Asp
			20					25					30		
Thr	Ala	Ser	Ala	Leu	Tyr	Arg	Glu	Ala	Leu	Glu	Ser	Pro	Glu	His	Cys
			35				40					45			
Ser	Pro	His	His	Thr	Ala	Leu	Arg	Gln	Ala	Ile	Leu	Cys	Trp	Gly	Glu
			50				55				60				
Leu	Met	Thr	Leu	Ala	Thr	Trp	Val	Gly	Val	Asn	Leu	Glu	Asp	Pro	Ala
65					70				75						80

Ser Arg Asp Leu Val Val Ser Tyr Val Asn Thr Asn Met Gly Leu Lys
 85 90 95
 Phe Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
 100 105 110
 Glu Thr Val Ile Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
 115 120 125
 Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
 130 135 140
 Glu Thr Thr Val Val Arg Arg Gly Arg Ser Pro Arg Arg Arg Thr
 145 150 155 160
 Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg Ser
 165 170 175
 Gln Ser Arg Leu Gly Pro Leu Leu Val Leu Gln Ala Gly Phe Phe Leu
 180 185 190
 Leu Thr Arg Ile Leu Thr Ile Pro Gln Ser Leu Asp Ser Trp Trp Thr
 195 200 205
 Ser Leu Asn Phe Leu Gly Gly Thr Thr Val Cys Leu Gly Gln Asn Ser
 210 215 220
 Gln Ser Pro Thr Ser Asn His Ser Pro Thr Ser Cys Pro Pro Thr Cys
 225 230 235 240
 Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg Phe Ile Ile Phe Leu Phe
 245 250 255
 Ile Leu Leu Leu Cys Leu Ile Phe Leu Leu Val Leu Leu Asp Tyr Gln
 260 265 270
 Gly Met Leu Pro Val Cys Pro Leu Ile Pro Gly Ser Ser Thr Thr Ser
 275 280 285
 Thr Gly Pro Cys Arg Thr Cys Met Thr Thr Ala Gln Gly Thr Ser Met
 290 295 300
 Tyr Pro Ser Cys Cys Cys Thr Lys Pro Ser Asp Gly Asn Cys Thr Cys
 305 310 315 320
 Ile Pro Ile Pro Ser Ser Trp Ala Phe Gly Lys Phe Leu Trp Glu Trp
 325 330 335
 Ala Ser Ala Arg Phe Ser Trp Leu Ser Leu Leu Val Pro Phe Val Gln
 340 345 350
 Trp Phe Val Gly Leu Ser Pro Thr Val Trp Leu Ser Val Ile Trp Met
 355 360 365
 Met Trp Tyr Trp Gly Pro Ser Leu Tyr Ser Ile Leu Ser Pro Phe Leu
 370 375 380
 Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val Tyr Ile
 385 390 395

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 870 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATG	GAC	ATC	GAC	CCT	TAT	AAA	GAA	TTT	GGA	GCT	ACT	GTG	GAG	TTA	CTC	48
Met	Asp	Ile	Asp	Pro	Tyr	Lys	Glu	Phe	Gly	Ala	Thr	Val	Glu	Leu	Leu	
1				5					10					15		
TCG	TTT	TTG	CCT	TCT	GAC	TTC	TTT	CCT	TCA	GTA	CGA	GAT	CTT	CTA	GAT	96
Ser	Phe	Leu	Pro	Ser	Asp	Phe	Phe	Pro	Ser	Val	Arg	Asp	Leu	Leu	Asp	
			20					25					30			

ACC	GCC	TCA	GCT	CTG	TAT	CGG	GAA	GCC	TTA	GAG	TCT	CCT	GAG	CAT	TGT	144
Thr	Ala	Ser	Ala	Leu	Tyr	Arg	Glu	Ala	Leu	Glu	Ser	Pro	Glu	His	Cys	
		35					40					45				
TCA	CCT	CAC	CAT	ACT	GCA	CTC	AGG	CAA	GCA	ATT	CTT	TGC	TGG	GGG	GAA	192
Ser	Pro	His	His	Thr	Ala	Leu	Arg	Gln	Ala	Ile	Leu	Cys	Trp	Gly	Glu	
	50					55					60					
CTA	ATG	ACT	CTA	GCT	ACC	TGG	GTG	GGT	GTT	AAT	TTG	GAA	GAT	CCA	GCG	240
Leu	Met	Thr	Leu	Ala	Thr	Trp	Val	Gly	Val	Asn	Leu	Glu	Asp	Pro	Ala	
	65				70					75					80	
TCT	AGA	GAC	CTA	GTA	GTC	AGT	TAT	GTC	AAC	ACT	AAT	ATG	GGC	CTA	AAG	288
Ser	Arg	Asp	Leu	Val	Val	Ser	Tyr	Val	Asn	Thr	Asn	Met	Gly	Leu	Lys	
				85					90					95		
TTC	AGG	CAA	CTC	TTG	TGG	TTT	CAC	ATT	TCT	TGT	CTC	ACT	TTT	GGA	AGA	336
Phe	Arg	Gln	Leu	Leu	Trp	Phe	His	Ile	Ser	Cys	Leu	Thr	Phe	Gly	Arg	
			100					105					110			
GAA	ACA	GTT	ATA	GAG	TAT	TTG	GTG	TCT	TTC	GGA	GTG	TGG	ATT	CGC	ACT	384
Glu	Thr	Val	Ile	Glu	Tyr	Leu	Val	Ser	Phe	Gly	Val	Trp	Ile	Arg	Thr	
		115					120					125				
CCT	CCA	GCT	TAT	AGA	CCA	CCA	AAT	GCC	CCT	ATC	CTA	TCA	ACA	CTT	CCG	432
Pro	Pro	Ala	Tyr	Arg	Pro	Pro	Asn	Ala	Pro	Ile	Leu	Ser	Thr	Leu	Pro	
	130					135					140					
GAG	ACT	ACT	GTT	GTT	AGA	CGA	CGA	GGC	AGG	TCC	CCT	AGA	AGA	AGA	ACT	480
Glu	Thr	Thr	Val	Val	Arg	Arg	Arg	Gly	Arg	Ser	Pro	Arg	Arg	Arg	Thr	
	145				150					155					160	
CCC	TCG	CCT	CGC	AGA	CGA	AGG	TCT	CAA	TCG	CCG	CGT	CGC	AGA	AGA	TCC	528
Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Pro	Arg	Arg	Arg	Arg	Ser	
				165					170					175		
TCA	ACA	ACC	AGC	ACG	GGA	CCA	TGC	CGG	ACC	TGC	ATG	ACT	ACT	GCT	CAA	576
Ser	Thr	Thr	Ser	Thr	Gly	Pro	Cys	Arg	Thr	Cys	Met	Thr	Thr	Ala	Gln	
			180					185					190			
GGA	ACC	TCT	ATG	TAT	CCC	TCC	TGT	TGC	TGT	ACC	AAA	CCT	TCG	GAC	GGA	624
Gly	Thr	Ser	Met	Tyr	Pro	Ser	Cys	Cys	Cys	Thr	Lys	Pro	Ser	Asp	Gly	
		195				200						205				
AAT	TGC	ACC	TGT	ATT	CCC	ATC	CCA	TCA	TCC	TGG	GCT	TTC	GGA	AAA	TTC	672
Asn	Cys	Thr	Cys	Ile	Pro	Ile	Pro	Ser	Ser	Trp	Ala	Phe	Gly	Lys	Phe	
	210					215					220					
CTA	TGG	GAG	TGG	GCC	TCA	GCC	CGT	TTC	TCC	TGG	CTC	AGT	TTA	CTA	GTG	720
Leu	Trp	Glu	Trp	Ala	Ser	Ala	Arg	Phe	Ser	Trp	Leu	Ser	Leu	Leu	Val	
	225				230					235					240	
CCA	TTT	GTT	CAG	TGG	TTC	GTA	GGG	CTT	TCC	CCC	ACT	GTT	TGG	CTT	TCA	768
Pro	Phe	Val	Gln	Trp	Phe	Val	Gly	Leu	Ser	Pro	Thr	Val	Trp	Leu	Ser	
			245					250						255		
GTT	ATA	TGG	ATG	ATG	TGG	TAT	TGG	GGG	CCA	AGT	CTG	TAC	AGC	ATC	TTG	816
Val	Ile	Trp	Met	Met	Trp	Tyr	Trp	Gly	Pro	Ser	Leu	Tyr	Ser	Ile	Leu	
			260					265					270			

AGT CCC TTT TTA CCG CTG TTA CCA ATT TTC TTT TGT CTT TGG GTA TAC
 Ser Pro Phe Leu Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val Tyr
 275 280 285

864

ATT TAA
 Ile

870

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 289 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met	Asp	Ile	Asp	Pro	Tyr	Lys	Glu	Phe	Gly	Ala	Thr	Val	Glu	Leu	Leu	
1				5					10					15		
Ser	Phe	Leu	Pro	Ser	Asp	Phe	Phe	Pro	Ser	Val	Arg	Asp	Leu	Leu	Asp	
			20					25					30			
Thr	Ala	Ser	Ala	Leu	Tyr	Arg	Glu	Ala	Leu	Glu	Ser	Pro	Glu	His	Cys	
		35					40					45				
Ser	Pro	His	His	Thr	Ala	Leu	Arg	Gln	Ala	Ile	Leu	Cys	Trp	Gly	Glu	
	50					55					60					
Leu	Met	Thr	Leu	Ala	Thr	Trp	Val	Gly	Val	Asn	Leu	Glu	Asp	Pro	Ala	
	65				70					75				80		
Ser	Arg	Asp	Leu	Val	Val	Ser	Tyr	Val	Asn	Thr	Asn	Met	Gly	Leu	Lys	
				85					90					95		
Phe	Arg	Gln	Leu	Leu	Trp	Phe	His	Ile	Ser	Cys	Leu	Thr	Phe	Gly	Arg	
			100					105					110			
Glu	Thr	Val	Ile	Glu	Tyr	Leu	Val	Ser	Phe	Gly	Val	Trp	Ile	Arg	Thr	
		115					120					125				
Pro	Pro	Ala	Tyr	Arg	Pro	Pro	Asn	Ala	Pro	Ile	Leu	Ser	Thr	Leu	Pro	
	130					135					140					
Glu	Thr	Thr	Val	Val	Arg	Arg	Gly	Arg	Ser	Pro	Arg	Arg	Arg	Arg	Thr	
	145				150				155					160		
Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Pro	Arg	Arg	Arg	Arg	Ser	
				165					170					175		
Ser	Thr	Thr	Ser	Thr	Gly	Pro	Cys	Arg	Thr	Cys	Met	Thr	Thr	Ala	Gln	
			180					185					190			
Gly	Thr	Ser	Met	Tyr	Pro	Ser	Cys	Cys	Thr	Lys	Pro	Ser	Asp	Gly		
		195					200				205					
Asn	Cys	Thr	Cys	Ile	Pro	Ile	Pro	Ser	Ser	Trp	Ala	Phe	Gly	Lys	Phe	
	210					215					220					
Leu	Trp	Glu	Trp	Ala	Ser	Ala	Arg	Phe	Ser	Trp	Leu	Ser	Leu	Leu	Val	
	225				230					235				240		
Pro	Phe	Val	Gln	Trp	Phe	Val	Gly	Leu	Ser	Pro	Thr	Val	Trp	Leu	Ser	
				245					250					255		
Val	Ile	Trp	Met	Met	Trp	Tyr	Trp	Gly	Pro	Ser	Leu	Tyr	Ser	Ile	Leu	
			260					265					270			
Ser	Pro	Phe	Leu	Pro	Leu	Leu	Pro	Ile	Phe	Phe	Cys	Leu	Trp	Val	Tyr	
		275					280					285				
Ile																

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1263 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATG GAT ATC AAT GCT TCT AGA GCC TTA GCC AAT GTG TAT GAT CTA CCA	48
Met Asp Ile Asn Ala Ser Arg Ala Leu Ala Asn Val Tyr Asp Leu Pro	
1 5 10 15	
GAT GAT TTC TTT CCA AAA ATA GAT GAT CTT GTT AGA GAT GCT AAA GAC	96
Asp Asp Phe Phe Pro Lys Ile Asp Asp Leu Val Arg Asp Ala Lys Asp	
20 25 30	
GCT TTA GAG CCT TAT TGG AAA TCA GAT TCA ATA AAG AAA CAT GTT TTG	144
Ala Leu Glu Pro Tyr Trp Lys Ser Asp Ser Ile Lys Lys His Val Leu	
35 40 45	
ATT GCA ACT CAC TTT GTG GAT CTT ATT GAA GAC TTC TGG CAG ACT ACA	192
Ile Ala Thr His Phe Val Asp Leu Ile Glu Asp Phe Trp Gln Thr Thr	
50 55 60	
CAG GGC ATG CAT GAA ATA GCC GAA TCA TTA AGA GCT GTT ATA CCT CCC	240
Gln Gly Met His Glu Ile Ala Glu Ser Leu Arg Ala Val Ile Pro Pro	
65 70 75 80	
ACT ACT ACT CCT GTT CCA CCG GGT TAT CTT ATT CAG CAC GAA GAA GCT	288
Thr Thr Thr Pro Val Pro Pro Gly Tyr Leu Ile Gln His Glu Glu Ala	
85 90 95	
GAA GAG ATA CCT TTG GGA GAT TTA TTT AAA CAC CAA GAA GAA AGG ATA	336
Glu Glu Ile Pro Leu Gly Asp Leu Phe Lys His Gln Glu Glu Arg Ile	
100 105 110	
GTG AGT TTC CAA CCC GAC TAT CCG ATT ACG GCT AGA ATT CAT GCT CAT	384
Val Ser Phe Gln Pro Asp Tyr Pro Ile Thr Ala Arg Ile His Ala His	
115 120 125	
TTG AAA GCT TAT GCA AAA ATT AAC GAG GAA TCA CTG GAT AGG GCT AGG	432
Leu Lys Ala Tyr Ala Lys Ile Asn Glu Glu Ser Leu Asp Arg Ala Arg	
130 135 140	
AGA TTG CTT TGG TGG CAT TAC AAC TGT TTA CTG TGG GGA GAA GCT CAA	480
Arg Leu Leu Trp Trp His Tyr Asn Cys Leu Leu Trp Gly Glu Ala Gln	
145 150 155 160	
GTT ACT AAC TAT ATT TCT CGC TTG CGT ACT TGG TTG TCA ACT CCT GAG	528
Val Thr Asn Tyr Ile Ser Arg Leu Arg Thr Trp Leu Ser Thr Pro Glu	
165 170 175	
AAA TAT AGA GGT AGA GAT GCC CCG ACC ATT GAA GCA ATC ACT AGA CCA	576
Lys Tyr Arg Gly Arg Asp Ala Pro Thr Ile Glu Ala Ile Thr Arg Pro	
180 185 190	

ATC Ile	CAG Gln	GTG Val	GCT Ala	CAG Gln	GGA Gly	GGC Gly	CGA Arg	AAA Lys	ACA Thr	ACT Thr	ACG Thr	GGT Gly	ACT Thr	AGA Arg	AAA Lys	624
		195					200					205				
CCT Pro	CGT Arg	GGA Gly	CTC Leu	GAA Glu	CCT Pro	AGA Arg	AGA Arg	AGA Lys	AAA Val	GTT Lys	AAA Thr	ACC Thr	ACA Thr	GTT Val	GTC Val	672
	210					215				220						
TAT Tyr	GGG Gly	AGA Arg	AGA Arg	CGT Arg	TCA Ser	AAG Lys	TCC Ser	CGG Arg	GGA Gly	AGG Arg	AGA Arg	GCC Ala	CCT Pro	ACA Thr	CCC Pro	720
	225				230					235					240	
CAA Gln	CGT Arg	GCG Ala	GGC Gly	TCC Ser	CCT Pro	CTC Leu	CCA Pro	CGT Arg	AGT Ser	TCG Ser	AGC Ser	AGC Ser	CAC His	CAT His	AGA Arg	768
				245					250					255		
TCC Ser	TTC Phe	GGG Gly	GGA Gly	ATA Ile	CTA Leu	GCT Ala	GGC Gly	CTA Leu	ATC Ile	GGA Gly	TTA Leu	CTG Leu	GTA Val	AGC Ser	TTT Phe	816
		260					265						270			
TTC Phe	TTG Leu	TTG Leu	ATA Ile	AAA Lys	ATT Ile	CTA Leu	GAA Glu	ATA Ile	CTG Leu	AGG Arg	AGG Arg	CTA Leu	GAT Asp	TGG Trp	TGG Trp	864
		275					280					285				
TGG Trp	ATT Ile	TCT Ser	CTC Leu	AGT Ser	TCT Ser	CCA Pro	AAG Lys	GGA Gly	AAA Lys	ATG Met	CAA Gln	TGC Cys	GCT Ala	TTC Phe	CAA Gln	912
	290					295					300					
GAT Asp	ACT Thr	GGA Gly	GCC Ala	CAA Gln	ATC Ile	TCT Ser	CCA Pro	CAT His	TAC Tyr	GTC Val	GGA Gly	TCT Ser	TGC Cys	CCG Pro	TGG Trp	960
	305				310					315					320	
GGA Gly	TGC Cys	CCA Pro	GGA Gly	TTT Phe	CTT Leu	TGG Trp	ACC Thr	TAT Tyr	CTC Leu	AGG Arg	CTT Leu	TTT Phe	ATC Ile	ATC Ile	TTC Phe	1008
				325					330				335			
CTC Leu	TTA Leu	ATC Ile	CTG Leu	CTA Leu	GTA Val	GCA Ala	GCA Ala	GGC Gly	TTG Leu	CTG Leu	TAT Tyr	CTG Leu	ACG Thr	GAC Asp	AAC Asn	1056
			340					345					350			
GGG Gly	TCT Ser	ACT Thr	ATT Ile	TTA Leu	GGA Gly	AAG Lys	CTC Leu	CAA Gln	TGG Trp	GCG Ala	TCG Ser	GTC Val	TCA Ser	GCC Ala	CTT Leu	1104
		355					360					365				
TTC Phe	TCC Ser	TCC Ser	ATC Ile	TCT Ser	TCA Ser	CTA Leu	CTG Leu	CCC Pro	TCG Ser	GAT Asp	CCG Pro	AAA Lys	TCT Ser	CTC Leu	GTC Val	1152
	370					375					380					
GCT Ala	TTA Leu	ACG Thr	TTT Phe	GGA Gly	CTT Leu	TCA Ser	CTT Leu	ATA Ile	TGG Trp	ATG Met	ACT Thr	TCC Ser	TCC Ser	TCT Ser	GCC Ala	1200
	385				390					395					400	
ACC Thr	CAA Gln	ACG Thr	CTC Leu	GTC Val	ACC Thr	TTA Leu	ACG Thr	CAA Gln	TTA Leu	GCC Ala	ACG Thr	CTG Leu	TCT Ser	GCT Ala	CTT Leu	1248
				405					410					415		
TTT Phe	TAC Tyr	AAG Lys	AGC Ser	TAG												1263
			420													

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 420 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met	Asp	Ile	Asn	Ala	Ser	Arg	Ala	Leu	Ala	Asn	Val	Tyr	Asp	Leu	Pro
1				5				10						15	
Asp	Asp	Phe	Phe	Pro	Lys	Ile	Asp	Asp	Leu	Val	Arg	Asp	Ala	Lys	Asp
			20					25					30		
Ala	Leu	Glu	Pro	Tyr	Trp	Lys	Ser	Asp	Ser	Ile	Lys	Lys	His	Val	Leu
		35					40					45			
Ile	Ala	Thr	His	Phe	Val	Asp	Leu	Ile	Glu	Asp	Phe	Trp	Gln	Thr	Thr
	50					55					60				
Gln	Gly	Met	His	Glu	Ile	Ala	Glu	Ser	Leu	Arg	Ala	Val	Ile	Pro	Pro
	65				70				75						80
Thr	Thr	Thr	Pro	Val	Pro	Pro	Gly	Tyr	Leu	Ile	Gln	His	Glu	Glu	Ala
			85						90					95	
Glu	Glu	Ile	Pro	Leu	Gly	Asp	Leu	Phe	Lys	His	Gln	Glu	Glu	Arg	Ile
			100					105					110		
Val	Ser	Phe	Gln	Pro	Asp	Tyr	Pro	Ile	Thr	Ala	Arg	Ile	His	Ala	His
		115					120					125			
Leu	Lys	Ala	Tyr	Ala	Lys	Ile	Asn	Glu	Glu	Ser	Leu	Asp	Arg	Ala	Arg
	130					135					140				
Arg	Leu	Leu	Trp	Trp	His	Tyr	Asn	Cys	Leu	Leu	Trp	Gly	Glu	Ala	Gln
	145				150				155						160
Val	Thr	Asn	Tyr	Ile	Ser	Arg	Leu	Arg	Thr	Trp	Leu	Ser	Thr	Pro	Glu
			165						170					175	
Lys	Tyr	Arg	Gly	Arg	Asp	Ala	Pro	Thr	Ile	Glu	Ala	Ile	Thr	Arg	Pro
			180					185					190		
Ile	Gln	Val	Ala	Gln	Gly	Gly	Arg	Lys	Thr	Thr	Thr	Gly	Thr	Arg	Lys
	195						200					205			
Pro	Arg	Gly	Leu	Glu	Pro	Arg	Arg	Lys	Val	Lys	Thr	Thr	Val	Val	
	210					215				220					
Tyr	Gly	Arg	Arg	Arg	Ser	Lys	Ser	Arg	Gly	Arg	Arg	Ala	Pro	Thr	Pro
	225					230				235					240
Gln	Arg	Ala	Gly	Ser	Pro	Leu	Pro	Arg	Ser	Ser	Ser	Ser	His	His	Arg
			245						250					255	
Ser	Phe	Gly	Gly	Ile	Leu	Ala	Gly	Leu	Ile	Gly	Leu	Leu	Val	Ser	Phe
		260					265						270		
Phe	Leu	Leu	Ile	Lys	Ile	Leu	Glu	Ile	Leu	Arg	Arg	Leu	Asp	Trp	Trp
		275					280					285			
Trp	Ile	Ser	Leu	Ser	Ser	Pro	Lys	Gly	Lys	Met	Gln	Cys	Ala	Phe	Gln
	290					295					300				
Asp	Thr	Gly	Ala	Gln	Ile	Ser	Pro	His	Tyr	Val	Gly	Ser	Cys	Pro	Trp
	305				310					315					320
Gly	Cys	Pro	Gly	Phe	Leu	Trp	Thr	Tyr	Leu	Arg	Leu	Phe	Ile	Ile	Phe
			325						330					335	
Leu	Leu	Ile	Leu	Leu	Val	Ala	Ala	Gly	Leu	Leu	Tyr	Leu	Thr	Asp	Asn
		340					345						350		
Gly	Ser	Thr	Ile	Leu	Gly	Lys	Leu	Gln	Trp	Ala	Ser	Val	Ser	Ala	Leu
		355					360					365			
Phe	Ser	Ser	Ile	Ser	Ser	Leu	Leu	Pro	Ser	Asp	Pro	Lys	Ser	Leu	Val
	370					375					380				
Ala	Leu	Thr	Phe	Gly	Leu	Ser	Leu	Ile	Trp	Met	Thr	Ser	Ser	Ser	Ala
	385				390				395						400

Thr Gln Thr Leu Val Thr Leu Thr Gln Leu Ala Thr Leu Ser Ala Leu
405 410 415
Phe Tyr Lys Ser
420

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 552 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATG GAC ATC GAC CCT TAT AAA GAA TTT GGA GCT ACT GTC GAG TTA CTC	48
Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu	
1 5 10 15	
TCG TTT TTG CCT TCT GAC TTC TTT CCT TCA GTA CGA GAT CTT CTA GAT	96
Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp	
20 25 30	
ACC GCC TCA GCT CTG TAT CGG GAA GCC TTA GAG TCT CCT GAG CAT TGT	144
Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys	
35 40 45	
TCA CCT CAC CAT ACT GCA CTC AGG CAA GCA ATT CTT TGC TGG GGG GAA	192
Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu	
50 55 60	
CTA ATG ACT CTA GCT ACC TGG GTG GGT GTT AAT TTG GAA GAT CCA GCG	240
Leu Met Thr Leu Ala Thr Trp Val Gly Val Asn Leu Glu Asp Pro Ala	
65 70 75 80	
TCT AGA GAC CTA GTA GTC AGT TAT GTC AAC ACT AAT ATG GGC CTA AAG	288
Ser Arg Asp Leu Val Val Ser Tyr Val Asn Thr Asn Met Gly Leu Lys	
85 90 95	
TTC AGG CAA CTC TTG TGG TTT CAC ATT TCT TGT CTC ACT TTT GGA ACA	336
Phe Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Thr	
100 105 110	
GAA ACA GTT ATA GAG TAT TTG GTG TCT TTC GGA GTG TGG ATT CGC ACT	384
Glu Thr Val Ile Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr	
115 120 125	
CCT CCA GCT TAT AGA CCA CCA AAT GCC OCT ATC CTA TCA ACA CTT CCG	432
Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro	
130 135 140	
GAG ACT ACT GTT GTT AGA CGA CCA GGC AGG TCC CCT AGA AGA AGA ACT	480
Glu Thr Thr Val Val Arg Arg Pro Gly Arg Ser Pro Arg Arg Arg Thr	
145 150 155 160	
CCC TCG CCT CGC AGA CGA AGG TCT CAA TCG CCC CGT CGC AGA AGA TCT	528
Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg Arg Ser	
165 170 175	

CAA TCT CGG GAA TCT CAA TGT TAG
 Gln Ser Arg Glu Ser Gln Cys
 180

552

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 183 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met	Asp	Ile	Asp	Pro	Tyr	Lys	Glu	Phe	Gly	Ala	Thr	Val	Glu	Leu	Leu	
1				5					10					15		
Ser	Phe	Leu	Pro	Ser	Asp	Phe	Phe	Pro	Ser	Val	Arg	Asp	Leu	Leu	Asp	
			20					25					30			
Thr	Ala	Ser	Ala	Leu	Tyr	Arg	Glu	Ala	Leu	Glu	Ser	Pro	Glu	His	Cys	
			35				40					45				
Ser	Pro	His	His	Thr	Ala	Leu	Arg	Gln	Ala	Ile	Leu	Cys	Trp	Gly	Glu	
	50					55					60					
Leu	Met	Thr	Leu	Ala	Thr	Trp	Val	Gly	Val	Asn	Leu	Glu	Asp	Pro	Ala	
	65				70				75					80		
Ser	Arg	Asp	Leu	Val	Ser	Tyr	Val	Asn	Thr	Asn	Met	Gly	Leu	Lys		
			85				90						95			
Phe	Arg	Gln	Leu	Leu	Trp	Phe	His	Ile	Ser	Cys	Leu	Thr	Phe	Gly	Thr	
			100				105						110			
Glu	Thr	Val	Ile	Glu	Tyr	Leu	Val	Ser	Phe	Gly	Val	Trp	Ile	Arg	Thr	
			115				120					125				
Pro	Pro	Ala	Tyr	Arg	Pro	Pro	Asn	Ala	Pro	Ile	Leu	Ser	Thr	Leu	Pro	
			130			135					140					
Glu	Thr	Thr	Val	Val	Arg	Arg	Pro	Gly	Arg	Ser	Pro	Arg	Arg	Arg	Thr	
					150				155					160		
Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Pro	Arg	Arg	Arg	Arg	Ser	
				165				170						175		
Gln	Ser	Arg	Glu	Ser	Gln	Cys										
															180	

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 681 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ATG	GAG	AAC	ATC	ACA	TCA	GGA	TTC	CTA	GGA	CCC	CTT	CTC	GTG	TTA	CAG	
Met	Glu	Asn	Ile	Thr	Ser	Gly	Phe	Leu	Gly	Pro	Leu	Leu	Val	Leu	Gln	
1				5				10						15		
GCG	GGG	TTT	TTC	TTG	TTG	ACA	AGA	ATC	CTC	ACA	ATA	CCC	CAG	AGT	CTA	
Ala	Gly	Phe	Phe	Leu	Leu	Thr	Arg	Ile	Leu	Thr	Ile	Pro	Gln	Ser	Leu	
			20				25					30				

48

96

GAC TCG TGG TGG ACT TCT CTC AAT TTT CTA GGG GGA ACT ACC GTG TGT Asp Ser Trp Trp Thr Ser Leu Asn Phe Leu Gly Gly Thr Thr Val Cys 35 40 45	144
CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC TCA CCA ACC TCT Leu Gly Gln Asn Ser Gln Ser Pro Thr Ser Asn His Ser Pro Thr Ser 50 55 60	192
TGT CCT CCA ACT TGT CCT GGT TAT CGC TGG ATG TGT CTG CGG CGT TTT Cys Pro Pro Thr Cys Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg Phe 65 70 75 80	240
ATC ATC TTC CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC TTG TTG GTT Ile Ile Phe Leu Phe Ile Leu Leu Leu Cys Leu Ile Phe Leu Leu Val 85 90 95	288
CTT CTG GAC TAT CAA GGT ATG TTG CCC GTT TGT CCT CTA ATT CCA GGA Leu Leu Asp Tyr Gln Gly Met Leu Pro Val Cys Pro Leu Ile Pro Gly 100 105 110	336
TCC TCA ACA ACC AGC ACG GGA CCA TGC CGG ACC TGC ATG ACT ACT GCT Ser Ser Thr Thr Ser Thr Gly Pro Cys Arg Thr Cys Met Thr Thr Ala 115 120 125	384
CAA GGA ACC TCT ATG TAT CCC TCC TGT TGC TGT ACC AAA CCT TCG GAC Gln Gly Thr Ser Met Tyr Pro Ser Cys Cys Cys Thr Lys Pro Ser Asp 130 135 140	432
GGA AAT TGC ACC TGT ATT CCC ATC CCA TCA TCC TGG GCT TTC GGA AAA Gly Asn Cys Thr Cys Ile Pro Ile Pro Ser Ser Trp Ala Phe Gly Lys 145 150 155 160	480
TTC CTA TGG GAG TGG GCC TCA GCC CCT TTC TCC TGG CTC AGT TTA CTA Phe Leu Trp Glu Trp Ala Ser Ala Pro Phe Ser Trp Leu Ser Leu Leu 165 170 175	528
GTC CCA TTT GTT CAG TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT Val Pro Phe Val Gln Trp Phe Val Gly Leu Ser Pro Thr Val Trp Leu 180 185 190	576
TCA GTT ATA TGG ATG ATG TGG TAT TGG GGG CCA AGT CTG TAC AGC ATC Ser Val Ile Trp Met Met Trp Tyr Trp Gly Pro Ser Leu Tyr Ser Ile 195 200 205	624
TTG AGT CCC TTT TTA CCG CTG TTA CCA ATT TTC TTT TGT CTT TGG GTA Leu Ser Pro Phe Leu Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val 210 215 220	672
TAC ATT TAA Tyr Ile 225	681

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 226 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met	Glu	Asn	Ile	Thr	Ser	Gly	Phe	Leu	Gly	Pro	Leu	Leu	Val	Leu	Gln
1				5					10					15	
Ala	Gly	Phe	Phe	Leu	Leu	Thr	Arg	Ile	Leu	Thr	Ile	Pro	Gln	Ser	Leu
			20					25					30		
Asp	Ser	Trp	Trp	Thr	Ser	Leu	Asn	Phe	Leu	Gly	Gly	Thr	Thr	Val	Cys
		35					40					45			
Leu	Gly	Gln	Asn	Ser	Gln	Ser	Pro	Thr	Ser	Asn	His	Ser	Pro	Thr	Ser
	50					55					60				
Cys	Pro	Pro	Thr	Cys	Pro	Gly	Tyr	Arg	Trp	Met	Cys	Leu	Arg	Arg	Phe
65					70					75					80
Ile	Ile	Phe	Leu	Phe	Ile	Leu	Leu	Leu	Cys	Leu	Ile	Phe	Leu	Leu	Val
				85					90					95	
Leu	Leu	Asp	Tyr	Gln	Gly	Met	Leu	Pro	Val	Cys	Pro	Leu	Ile	Pro	Gly
		100						105					110		
Ser	Ser	Thr	Thr	Ser	Thr	Gly	Pro	Cys	Arg	Thr	Cys	Met	Thr	Thr	Ala
		115					120					125			
Gln	Gly	Thr	Ser	Met	Tyr	Pro	Ser	Cys	Cys	Cys	Thr	Lys	Pro	Ser	Asp
	130					135					140				
Gly	Asn	Cys	Thr	Cys	Ile	Pro	Ile	Pro	Ser	Ser	Trp	Ala	Phe	Gly	Lys
145					150					155					160
Phe	Leu	Trp	Glu	Trp	Ala	Ser	Ala	Pro	Phe	Ser	Trp	Leu	Ser	Leu	Leu
			165					170						175	
Val	Pro	Phe	Val	Gln	Trp	Phe	Val	Gly	Leu	Ser	Pro	Thr	Val	Trp	Leu
			180					185					190		
Ser	Val	Ile	Trp	Met	Met	Trp	Tyr	Trp	Gly	Pro	Ser	Leu	Tyr	Ser	Ile
		195					200					205			
Leu	Ser	Pro	Phe	Leu	Pro	Leu	Leu	Pro	Ile	Phe	Phe	Cys	Leu	Trp	Val
	210					215					220				
Tyr	Ile														
225															